

#2

520.39527X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE



Applicant(s): SAEKI, ET AL.
Serial No.:
Filed: January 18, 2001
Title: MULTIDIMENSIONAL CROSSBAR NETWORK AND
PARALLEL COMPUTER SYSTEM
Group:

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

January 18, 2001

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2000-072458 filed March 10, 2000.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP

Carl I. Brundidge
Registration No. 29,621

CIB/mdt
Attachment
(703)312-6600

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2000年 3月10日

出 願 番 号
Application Number: 特願2000-072458

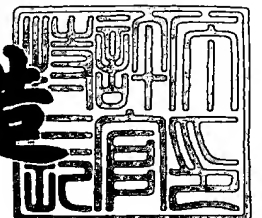
出 願 人
Applicant(s): 株式会社日立製作所



2000年10月20日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3086898

【書類名】 特許願

【整理番号】 PNT990954

【提出日】 平成12年 3月10日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/173

【発明者】

【住所又は居所】 神奈川県秦野市堀山下1番地 株式会社日立製作所 エ
ンタープライズサーバ事業部内

【氏名】 佐伯 裕治

【発明者】

【住所又は居所】 神奈川県秦野市堀山下1番地 株式会社日立製作所 エ
ンタープライズサーバ事業部内

【氏名】 中畑 昌也

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 多次元クロスバーネットワークおよび並列計算機システム

【特許請求の範囲】

【請求項 1】

論理的に多次元配列された複数の演算ノードを複数のクロスバースイッチによって相互接続した多次元クロスバーネットワークにおいて、

第 1 のクロスバースイッチと第 2 のクロスバースイッチとを接続する各パケット伝送路上に、上記第 1、第 2 のクロスバースイッチと第 3 のクロスバースイッチに接続されるスイッチング装置を備え、

上記スイッチング装置によって、上記第 1、第 2、第 3 のクロスバースイッチ間のパケット交換を行うと共に、上記何れかのクロスバースイッチと光信号でパケット通信するためのインタフェース変換を行うことを特徴とする多次元クロスバーネットワーク。

【請求項 2】

論理的に多次元配列された複数の演算ノードを相互接続するための複数の X、Y、Z 軸クロスバースイッチからなる多次元クロスバーネットワークにおいて、

X 軸クロスバースイッチと Y 軸クロスバースイッチとの間の各パケット伝送路上に、上記 X 軸および Y 軸クロスバースイッチからの受信パケットを Z 軸クロスバースイッチに選択的に転送するためのスイッチング装置を有することを特徴とする多次元クロスバーネットワーク。

【請求項 3】

前記スイッチング装置が、前記 X 軸クロスバースイッチの入出力ポートのうちの 1 つに接続するための第 1 入出力ポートと、前記 Y 軸クロスバースイッチの入出力ポートのうちの 1 つに接続するための第 2 入出力ポートと、発光素子と受光素子からなる光モジュールに接続するための第 3 入出力ポートと、上記第 1、第 2 入出力ポートと上記第 3 入出力ポートの間で送受信パケットのインタフェース変換を行うための手段とを備え、

前記 Y 軸クロスバースイッチと Z 軸クロスバースイッチとの間のパケット伝送路の一部が上記光モジュールに結合される光ファイバによって構成されることを

特徴とする請求項 2 に記載の多次元クロスバーネットワーク。

【請求項 4】

前記インタフェース変換手段が、前記第 1、第 2 入出力ポートで送受信されるパケットと前記第 3 入出力ポートで送受信されるパケットとを互いに独立した異なる同期クロックによって処理することを特徴とする請求項 2 または請求項 3 に記載の多次元クロスバーネットワーク。

【請求項 5】

それぞれ多次元配列された複数の演算ノードを複数の X、Y、Z 軸クロスバースイッチによって相互接続した第 1、第 2 のクロスバーネットワークからなる多次元クロスバーネットワークにおいて、

上記各クロスバーネットワークが、それぞれ 3 次元座標系で同一の Y、Z 座標値をもつ複数の演算ノード間で X 軸方向のパケット交換を行う複数の X 軸クロスバースイッチと、3 次元座標系で同一 Z 軸座標値をもつ演算ノードを収容している複数の X 軸クロスバースイッチ間で Y 軸方向のパケット交換を行う複数の Y 軸クロスバースイッチ群と、上記複数の Y 軸クロスバースイッチ群の間で Z 軸方向のパケット交換を行う複数の Z 軸クロスバースイッチとからなり、

上記第 1、第 2 のクロスバーネットワークにおける互いに対応した位置関係にある 2 つの Y 軸クロスバースイッチを該 Y 軸クロスバースイッチと Z 軸クロスバースイッチとの間の各パケット経路上に配置された複数のスイッチング L S I によって結合し、上記各スイッチング L S I によって上記第 1、第 2 のクロスバーネットワーク間のパケット交換を行うようにしたことを特徴とする多次元クロスバーネットワーク。

【請求項 6】

前記各スイッチング L S I が、前記 2 つの Y 軸クロスバースイッチに接続するための第 1、第 2 入出力ポートと、発光素子と受光素子からなる第 1、第 2 の光モジュールに接続するための第 3、第 4 の入出力ポートと、上記第 1、第 2 の入出力ポートからの受信パケットをヘッダ情報に従って上記第 1、第 2 の入出力ポートのうちの他方、上記第 3 または第 4 の入出力ポートに選択的に出力すると共に、上記第 3、第 4 の入出力ポートからの受信パケットをそれぞれ上記第 1、第

2の入出力ポートに転送するための手段とからなり、

前記各Y軸クロスバースイッチとZ軸クロスバースイッチとの間のパケット伝送路の一部を上記各光モジュールに結合される光ファイバによって構成したことを特徴とする請求項5に記載の多次元クロスバーネットワーク。

【請求項7】

それぞれ3次元座標系で同一のY、Z座標値をもつ複数の演算ノード間でX軸方向のパケット交換を行うための複数のX軸クロスバースイッチと、それぞれ3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行うための複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイッチング手段とからなる多次元クロスバーネットワークにおいて、上記各Z軸クロスバースイッチング手段が、

3次元座標系で同一のX軸座標値をもつ複数のY軸クロスバースイッチの互いに対応したX軸座標位置にある入出力ポートに接続される第1入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第2入出力ポート群とを有する第1のスイッチングLSIと、

3次元座標系で同一のX軸座標値をもつ他の複数のY軸クロスバースイッチの互いに対応したX軸座標位置にある入出力ポートに接続される第1入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第2入出力ポート群とを有する第2のスイッチングLSIと、

上記第1のスイッチングLSIの第2入出力ポート群と上記第2のスイッチングLSIの第2入出力ポート群との間にそれぞれ光モジュールを介して結合される複数対の光ファイバとからなることを特徴とする多次元クロスバーネットワーク。

【請求項8】

それぞれL個の演算ノード収容し、L個の外部入出力ポートをもつX軸クロスバースイッチが搭載された複数のX軸ボードと、

M個の入出力ポートをもつY軸クロスバースイッチと、上記X軸クロスバースイッチに接続するための外部入出力ポートと光信号送受信用光モジュールに接続

するための入出力ポートを有し、上記 Y 軸クロスバースイッチの各入出力ポートに接続される複数のスイッチング L S I とが搭載された複数の Y 軸ボードと、

N 個の入出力ポートをもつ Z 軸クロスバースイッチと、光信号送受信用光モジュールに接続するための入出力ポートを有し、上記 Z 軸クロスバースイッチの各入出力ポートに接続される複数のインタフェース変換 L S I とが搭載された複数の Z 軸ボードと、

上記 X 軸クロスバースイッチの外部入出力ポートと上記 Y 軸ボード上の外部入出力ポートとの間を接続する複数の電気信号線と、

上記 Y 軸ボード上の光信号送受信用光モジュールと上記 Z 軸ボード上の光信号送受信用光モジュールとの間を接続する光信号線と
からなることを特徴とした $L \times M \times N$ の多次元配列並列計算機システム。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、多次元クロスバーネットワークおよび多次元配列並列計算機システムに関し、更に詳しくは、多数の演算装置間を接続するためのネットワーク構成に関する。

【 0 0 0 2 】

【従来の技術】

膨大な計算処理を高速に実行するために、例えば、複数の演算装置を密結合したマルチプロセッサ型計算機を 1 つの演算ノードとして、多数のノード間をネットワークで相互接続した構成の並列計算機システムが知られている。

並列計算機システムでは、多次元配列された各演算ノード毎にメモリを配置しておき、各演算ノードが、計算に必要なデータの一部を他ノードのメモリから自ノード内のメモリにコピーする分散メモリ型のものが多い。

上記分散メモリ型の並列計算機上で動作させるためにメモリ共有モデルから書き換えられたプログラムは、全ノードが同一プログラム／異データで動作するモデルに該当し、演算動作とノード間通信の現れ方が、例えば、通信中の各ノード対の一方のノードからみて、相手ノードの相対位置と通信データ量とが各ノード

対で一定となる等方転送になることが多い。メモリ共有モデルにおける最も基本的なデータアクセスは、メモリ空間の連続したアドレスでデータをアクセスするパターンを有し、上記基本パターンを分散メモリに素直に割り当てると、隣接ノード間でのデータ転送（隣接転送）が最も頻繁に現れる。

【0003】

この種の並列計算機システムでは、任意の2つのノード間で通信できるように演算ノードネットワークを構成する必要がある。現在では、例えば、メッシュ型、トーラス型、多次元クロスバー型、完全クロスバー型などと呼ばれるネットワーク構成によって、数百台～数千台の演算ノードからなる並列計算機システムが実現可能となっている。

「メッシュ型」および「トーラス型」のネットワーク構成では、各演算ノードに多次元論理座標を割り当てたとき、座標空間で互いに隣接した位置関係にある演算ノード間にのみネットワーク配線が存在する。従って、多数の演算ノードを論理座標と対応がとれた形で物理的に配置できれば、短い配線長で非常に多数の演算ノードを接続できる利点がある。しかしながら、このタイプのネットワーク構成では、上述した隣接転送は経路競合なしに行えるが、非隣接ノードのデータを頻繁にアクセスする等方転送の計算処理では、ノード間通信中に経路競合が発生し、システムの実効性能が低下する。従って、メッシュ型、トーラス型構成の並列計算機は、特定用途の大規模計算に適しており、汎用性には欠けるという性質がある。

【0004】

「完全クロスバー型」のネットワーク構成は、任意に組み合わされた複数対のノード間で同時にパケット通信できるため、各演算ノードに分散して存在するデータをグローバルにアクセスする等方転送の計算処理でも、ノード間通信ロスが少なく、効率的に計算処理を実行できる。完全クロスバー型のネットワークでは、クロスバースイッチを構成するLSI群を一箇所に集中して配置した場合、各演算ノードとクロスバースイッチとの接続配線長が長距離化する。

【0005】

「多次元クロスバー型」のネットワークは、上記2つの形式の中間的な論理構

造をもつもので、ノード間のグローバルな通信を比較的効率的に実行でき、接続ノード数とシステム性能のスケラブルな拡張を可能とする。

多次元クロスバーネットワークは、例えば、次元数が3の場合、 $L \times M \times N$ の直方体の各格子点に演算ノードをマップしておき、X軸方向に並ぶL個のノード間を接続するクロスバースイッチと、Y軸方向に並ぶM個のノード間を接続するクロスバースイッチと、Z軸方向に並ぶN個のノード間を接続するクロスバースイッチとによって多段クロスバースイッチを構成し、演算ノード毎に、ノード内メモリと、各ノードを上記X、Y、Z軸方向の各クロスバーに選択的に接続するためのExchangerと呼ばれるスイッチとを設けた論理構成となっている。

しかしながら、多次元クロスバーネットワークでは、各ノード毎に設けられるExchangerに、X、Y、Z軸方向の各クロスバースイッチと選択的に接続するための3系統の配線が必要となるため、何れかの系統が長距離集中配線となってしまふ。

【0006】

例えば、X軸クロスバースイッチと該X軸クロスバースイッチに接続されるL個の演算ノードを同一の基板上に搭載し、それぞれ同一Z座標のノード群を含むM枚の基板を同一のバックボードに収容し、該バックボード上に搭載されたL個のY軸クロスバーによって、同一X座標をもつM個のノード間を接続することによって、2次元配列のノード群までは比較的近接した配線によってネットワーク化できる。N個の2次元ノード配列からなる $L \times M \times N$ の3次元ノード配列は、N枚のバックボードに分散配置された合計 $L \times M \times N$ 個のノードを $L \times M$ 系統のZ軸クロスバースイッチの何れかのポートに接続することによって得られる。

【0007】

多次元クロスバー型は、完全クロスバー型に比較して、各軸方向のクロスバースイッチを構成するLSIの接続ポート数が少なく済むため、設計が容易になる。また、クロスバースイッチ用のLSI群と演算ノード群を幾つかのボードに分散配置し、各ボード上で部分的に配線を済ませることによって、完全クロスバー型のような演算ノードからクロスバースイッチ部への配線の集中を回避できる。但し、上述した $L \times M$ 系統のZ軸クロスバースイッチ群を一箇所に集めて配置

した場合、バックボード上の各演算ノードとZ軸クロスバースイッチとの接続に長距離配線が必要となる。

【0008】

上記多次元クロスバー型を改良したネットワーク構成として、X、Y、Zの各軸方向への配線源となるExchanger機能を分割し、演算ノードとX軸クロスバースイッチとの間の乗換え機能は各演算ノードにもたせ、X軸クロスバースイッチとY軸クロスバースイッチとの間の乗換え機能はX軸クロスバースイッチに、Y軸クロスバースイッチとZ軸クロスバースイッチとの間の乗換え機能はY軸クロスバースイッチにもたせた構成が「分散Exchanger方式」の多次元クロスバーネットワークとして知られている。分散Exchanger方式によれば、演算ノードからX軸、Y軸、Z軸クロスバースイッチと順次に辿って、配線を一箇所に収束する接続形態でネットワークを構成できる。

【0009】

【発明が解決しようとする課題】

しかしながら、分散Exchanger方式の多次元クロスバーネットワーク構成では、Y軸クロスバースイッチとZ軸クロスバースイッチとの間の接続配線が、システム規模に応じて長距離化するため、接続ケーブルの電気的特性によって、接続可能な演算ノード数やシステム性能が制限される。

また、分散Exchanger方式の多次元クロスバーネットワークでは、最上位軸（3次元の場合はZ軸）以外の各軸方向のクロスバースイッチにおいて、別軸方向のクロスバースイッチへの乗換え機能をもたせるために、各演算ノードがExchanger機能をもつ通常の多次元クロスバーネットワークのクロスバースイッチに比較して、約2倍の外部接続ピン（LSIピン）が必要となる。

【0010】

例えば、通常の多次元クロスバーネットワーク構成では、 $L \times M \times N$ 個のノードを接続する場合、Y軸クロスバースイッチには、Y軸方向のノード数に等しいMセットの入出力ポートがあればよい。これに対して、分散Exchanger方式のY軸クロスバースイッチには、X軸クロスバースイッチに接続するためのMセットの入出力ポートと、Z軸クロスバースイッチに接続するためのMセットの入出力

ポートが必要となるため、LSIの外部接続ピン数を同一にした場合、Y軸方向に配列できるノード数が半分になってしまう。

また、分散Exchanger方式のクロスバーネットワークでは、Z軸方向に隣接する2つのノード間で通信する場合でも、X軸クロスバーとY軸クロスバーを経由しなければZ軸クロスバースイッチに到達できない。従って、最も頻繁に現れる隣接転送のノード間通信において、隣接ノードの位置関係によっては、送信パケットが宛先ノードに到達する迄の所要時間が長くなるという問題がある。

【0011】

本発明の目的は、ノード間接続コストの上昇を抑えてシステム規模（接続可能ノード数）を拡張できる分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムを提供することにある。

本発明の他の目的は、長距離配線区間における特性劣化の少ない分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムを提供することにある。

本発明の更に他の目的は、クロスバースイッチ当たりの収容ノード数を増加できる分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムを提供することにある。

【0012】

【課題を解決するための手段】

上記目的を達成するために、本発明は、論理的に多次元配列された複数の演算ノードを複数のクロスバースイッチによって相互接続した多次元クロスバーネットワークにおいて、第1のクロスバースイッチと第2のクロスバースイッチとを接続するパケット伝送路上に、上記第1、第2のクロスバースイッチと第3のクロスバースイッチに接続されるスイッチング装置を備え、上記スイッチング装置によって、上記第1、第2、第3のクロスバースイッチ間のパケット交換を行うと共に、上記何れかのクロスバースイッチと光信号でパケット通信するためのインタフェース変換を行うことを特徴とする。

本発明によれば、クロスバースイッチ間を接続するスイッチング装置（LSI）に、光信号でパケット通信するためのインタフェース変換機能をもたせたこと

によって、距離の長い配線区間では光信号によってパケット通信できる。この場合、電気信号ケーブル接続ポートで送受信されるパケットと、光信号ケーブル（光ファイバ）接続ポートで送受信されるパケットとを互いに独立した異なる同期クロックによって処理することによって、長距離配線区間で光信号伝送の適した通信形態を採用できる。

【 0 0 1 3 】

本発明において、上記スイッチング装置は、例えば、座標軸の異なるクロスバースイッチ間で受信パケットを他の座標軸のクロスバースイッチに転送するパケット分岐用スイッチとして適用される。本発明の 1 実施例では、論理的に多次元配列された演算ノード群を相互接続するための複数の X、Y、Z 軸クロスバースイッチからなる多次元クロスバーネットワークおよび並列計算機システムにおいて、X 軸クロスバースイッチと Y 軸クロスバースイッチとの間の各パケット伝送路上に、上記 X 軸および Y 軸クロスバースイッチからの受信パケットを Z 軸クロスバースイッチに選択的に転送するためのスイッチング装置を有することを特徴とする。

上記構成によれば、Y 軸クロスバースイッチの外部で Z 軸クロスバースイッチへの乗り換えが実行されるため、Y 軸クロスバースイッチ L S I には、Z 軸クロスバースイッチとの接続ピンが不要となり、これらの接続ピンを X 軸クロスバースイッチとの接続用として有効に利用できる。また、Z 軸方向に隣接するノード間で通信を行う場合に、X 軸クロスバースイッチから Y 軸クロスバースイッチを経由せずに Z 軸クロスバーにパケットを転送できるため、通信パケットが宛先ノードに到達する迄の所要時間を短縮できる。

【 0 0 1 4 】

本発明において、上記スイッチング装置は、例えば、異なる多次元クロスバーネットワーク間でのパケット交換用スイッチとしても適用される。本発明の第 2 実施例では、それぞれ多次元配列された複数の演算ノードを複数の X、Y、Z 軸クロスバースイッチによって相互接続した第 1、第 2 のクロスバーネットワークからなる多次元クロスバーネットワークにおいて、上記各クロスバーネットワークが、それぞれ 3 次元座標系で同一の Y、Z 座標値をもつ複数の演算ノード間で

X軸方向のパケット交換を行う複数のX軸クロスバースイッチと、3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行う複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイッチとからなり、上記第1、第2のクロスバーネットワークにおける互に対応した位置関係にある2つのY軸クロスバースイッチを該Y軸クロスバースイッチとZ軸クロスバースイッチとの間の各パケット経路上に配置された複数のスイッチングLSIによって結合し、上記各スイッチングLSIによって上記第1、第2のクロスバーネットワーク間のパケット交換を行うようにしたことを特徴とする。

【0015】

この場合、スイッチングLSIは、具体的には、上記2つのY軸クロスバースイッチに接続するための第1、第2入出力ポートと、発光素子と受光素子からなる第1、第2の光モジュールに接続するための第3、第4の入出力ポートと、上記第1、第2の入出力ポートからの受信パケットをヘッダ情報に従って上記第1、第2の入出力ポートのうちの他方、上記第3または第4の入出力ポートに選択的に出力すると共に、上記第3、第4の入出力ポートからの受信パケットをそれぞれ上記第1、第2の入出力ポートに転送するための手段とからなり、上記各Y軸クロスバースイッチとZ軸クロスバースイッチとの間のパケット伝送路の一部が上記各光モジュールに結合される光ファイバによって構成される。

【0016】

本発明において、上記スイッチング装置は、例えば、多次元クロスバーネットワークにおける特定軸のクロスバースイッチに代わるパケット交換用スイッチとしても適用される。本発明の第3実施例では、それぞれ3次元座標系で同一のY、Z座標値をもつ複数の演算ノード間でX軸方向のパケット交換を行うための複数のX軸クロスバースイッチと、それぞれ3次元座標系で同一Z軸座標値をもつ演算ノードを収容している複数のX軸クロスバースイッチ間でY軸方向のパケット交換を行うための複数のY軸クロスバースイッチ群と、上記複数のY軸クロスバースイッチ群の間でZ軸方向のパケット交換を行う複数のZ軸クロスバースイ

ツチング手段とからなる多次元クロスバーネットワークにおいて、上記各 Z 軸クロスバースイッチング手段が、3 次元座標系で同一の X 軸座標値をもつ複数の Y 軸クロスバースイッチの互いに対応した X 軸座標位置にある入出力ポートに接続される第 1 入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第 2 入出力ポート群とを有する第 1 のスイッチング L S I と、3 次元座標系で同一の X 軸座標値をもつ他の複数の Y 軸クロスバースイッチの互いに対応した X 軸座標位置にある入出力ポートに接続される第 1 入出力ポート群と、それぞれ発光素子と受光素子とを含む複数の光モジュールに接続される第 2 入出力ポート群とを有する第 2 のスイッチング L S I と、上記第 1 のスイッチング L S I の第 2 入出力ポート群と上記第 2 のスイッチング L S I の第 2 入出力ポート群との間にそれぞれ光モジュールを介して結合される複数対の光ファイバとからなることを特徴とする。上記構成によれば、ノード間接続コストを抑えて 3 次元クロスバーネットワークを構成できる。

【 0 0 1 7 】

【発明の実施の形態】

以下、本発明の実施例を図面を参照して説明する。

図 1 は、本発明による分散 Exchanger 方式の 3 次元クロスバーネットワークにおける X、Y、Z 軸方向のクロスバースイッチ間の接続関係を説明するための図である。

図において、20 は X 軸クロスバースイッチ (L S I) 21 を搭載した X 軸ボード、30 は Y 軸クロスバースイッチ (L S I) 31 を搭載した Y 軸ボード、40 は Z 軸クロスバースイッチ (L S I) 41 を搭載した Z 軸ボードを示す。

【 0 0 1 8 】

L×M×N 個のノードを接続する 3 次元クロスバーネットワークを構成する場合、X 軸ボード 20 には、X 軸上の座標値と対応する L 個のノードボード 10 が収容される。22-1～22-L は、これらのノードボードを接続するためのコネクタを示す。各コネクタ 22 は、X 軸ボードの表面あるいは内部に形成されたプリント配線によって X 軸クロスバースイッチ 21 の入出力ポートと結合されている。

【 0 0 1 9 】

各ノードボード10は、図2に示すように、複数のLSIからなる演算処理部11と、ネットワークインタフェースLSI:12と、これらのLSIで共有して使用されるメモリLSI:13と、メモリスイッチ14とを有し、これらの回路要素によって演算ノードが形成されている。15は、ボード端部に設けられたリード端子群であり、これらのリード端子をX軸ボード20上のコネクタ22-iに挿し込むことによって、演算ノードがX軸クロスバースイッチ21の所定の入出力ポートに接続される。

演算処理部11は、他の演算ノードにデータを送信するとき、メモリ13の所定の領域に送信データと制御情報を設定した後、ネットワークインタフェースLSI:12を起動する。起動されたネットワークインタフェースLSI:12は、メモリ13に用意された制御情報に従って送信データを読み出し、例えば、8バイト幅の packets を生成して、X軸クロスバースイッチ21に送出する。上記 packets のヘッダ部には、宛先アドレスとして、3次元クロスバーネットワークにおける宛先ノード位置を示すX、Y、Z座標値が設定される。

【 0 0 2 0 】

以下、宛先アドレス[i, j, k]をもつ packets に着目して、図1の3次元クロスバーネットワーク上での転送動作を説明する。

X軸クロスバースイッチ21は、上述したノードボード(演算ノード)10を収容するためのL個の入出力ポートの他に、Y軸クロスバースイッチ31との接続線を収容するためのL個の外部入出力ポートを備える。X軸クロスバースイッチ21とY軸クロスバースイッチ31との接続には、例えば、同軸ケーブルからなる入出力線200[1, y, z]~200[L, y, z]が利用される。

【 0 0 2 1 】

ここで、鍵括弧内に文字は、3次元クロスバーネットワークにおける演算ノードのアドレス(X、Y、Z座標値)を示しており、例えば、入出力線200[1, y, z]には、コネクタ22-1に接続されたアドレス[1, y, z]をもつ演算ノードから出力された他の演算ノード宛の送信 packets 、あるいは他の演算ノードから上記アドレス[1, y, z]をもつ演算ノード宛の受信 packets が流

れることを意味している。また、座標値 $[y, z]$ は、X軸クロスバースイッチ 2 1 に收容された演算ノード群に共通するY、Z座標値であり、X軸クロスバースイッチ 2 1 には、上記座標値 $[y, z]$ がボードアドレスとして予め設定されている。

【0 0 2 2】

X軸クロスバースイッチ 2 1 は、收容している何れかの演算ノードから宛先アドレス $[i, j, k]$ のパケットを受信すると、パケットヘッダの宛先X座標に従って、第 i ポートにスイッチングする。また、受信パケットの宛先Y、Z座標 j, k を上記X軸ボードアドレスと比較し、Y軸上またはZ軸上でのスイッチングを要する受信パケットについては、上記第 i ポートの外部入出力線、例えば、2 0 0 $[i, y, z]$ に送出する (i = 宛先X座標値)。宛先Y、Z座標値 j, k がボードアドレス $[y, z]$ と一致し、Y軸、Z軸上でのスイッチングを必要としないパケットについては、第 i ポートのノードボード側 (コネクタ 2 2 - i) に送出する。

【0 0 2 3】

Y軸ボード 3 0 は、図 3、図 4 で後述するように、3次元クロスバーネットワークのZ座標値毎に、X座標と対応してL個ずつ用意され、それぞれボードアドレスとして特定のX、Z座標値が割り当てられている。各Y軸ボードには、Y軸クロスバースイッチ 3 1 と、3次元クロスバーネットワークのY座標と対応したM個のスイッチングLSI : 3 1 - 1 ~ 3 1 - Mとが搭載されている。

本実施例において、各スイッチングLSI : 3 1 は、X軸、Y軸のクロスバースイッチ間でZ軸クロスバースイッチ方向にパケットを分岐し、分岐パケットについて光伝送のためのインタフェース変換を行うためのものであり、以下の説明では、インタフェース変換LSIと呼ぶことにする。

【0 0 2 4】

X軸クロスバースイッチ 2 1 から入出力線 2 0 0 $[i, y, z]$ に送出されたパケットは、X、Z座標値 $[i, z]$ をボードアドレスとするY軸ボード 3 0 の第 y 番目のインタフェース変換LSI : 3 1 - y に入力される。上記インタフェース変換LSI : 3 1 - y は、入出力配線 2 0 0 $[i, y, z]$ からの受信パケ

ットの宛先アドレスをチェックし、Y軸上でのスイッチングを必要としないパケット、即ち、宛先Y座標値jが上記インタフェース変換LSI: 31-yのY座標yに一致したパケットは、レーザー発光素子33T-yと受光素子33R-yとからなる光モジュール33-yに接続された入出力線320-y (= 320-j) に送出し、Y軸上でのスイッチングを要するパケット、すなわち、宛先Y座標値jがインタフェース変換LSIのY座標値yに一致しないパケットは、Y軸クロスバースイッチ31に接続された入出力線310-yに出力する。

【0025】

入出力線310-yに出力されたパケットは、Y軸クロスバースイッチ31によって、宛先Y座標値jと対応した第j番目のインタフェース変換LSI: 31-jにスイッチングされる。上記インタフェース変換LSI: 31-jは、Y軸クロスバースイッチ31からの受信パケットについて宛先アドレスをチェックし、Z軸上でのスイッチングを必要としないパケット、すなわち、宛先Z座標値kがY軸ボードアドレスのZ座標値zと一致したパケットについては、入出力線200[i, j, z]に出力し、Z軸上でのスイッチングを要するパケットは、入出力回線320-jに出力する。

入出力線320-jに出力されたパケットは、レーザー発光素子33T-jで光信号に変換され後、入出力線（光ファイバ）300[i, j, z]に送出される。上記入出力線（光ファイバ）300[i, j, z]は、ボードアドレスとしてX、Y座標値[i, j]をもつZ軸ボード40に接続されている。

【0026】

各Z軸ボード40は、Z軸クロスバースイッチ41と、3次元クロスバーネットワークのZ座標値と対応したN個のインタフェース変換LSI: 42-1~42-Nとを有し、各インタフェース変換LSI: 42-q (q=1~N)は、Z軸クロスバースイッチ41に接続された入出力線410-qと、レーザー発光素子43T-qと受光素子43R-qとを含む光モジュール43-qに接続された入出力線420-qとを備えている。

【0027】

Y軸ボード30上のインタフェース変換LSI: 32とは異なり、Z軸ボード

4 0 上の各インタフェース変換 L S I : 4 2 の入出力ポートは、Z 軸クロスバースイッチ側と光モジュール側の 2 つとなっている。このため、Y 軸ボード 3 0 から入出力線（光ファイバ）3 0 0 [i , j , z] に光信号として送出されたパケットは、Z 軸ボード 4 0 側の受光素子 4 4 R - z で電気信号に変換された後、無条件で入出力線 4 1 0 - z に転送され、Z 軸クロスバースイッチ 4 1 に入力される。Z 軸クロスバースイッチ 4 1 は、上記パケットを宛先 Z 座標値 k に対応した第 k インタフェース変換 L S I : 4 2 - k にスイッチングする。

上記第 k インタフェース変換 L S I : 4 2 - k は、Z 軸クロスバースイッチ 4 1 からの入力パケットを入出力線 4 2 0 - k に転送する。従って、上記パケットは、レーザー発光素子 4 3 T - k で光信号に変換された後、光ファイバ 3 0 0 [i , j , k] に送出される。

【 0 0 2 8 】

上記光ファイバ 3 0 0 [i , j , k] は、ボードアドレスとして X、Z 座標値 [i , k] をもつ Y 軸ボード 3 0 の第 j インタフェース変換 L S I : 3 2 - j の光モジュールに接続されている。この時点では、受信パケットの宛先 X、Z 座標は、上記 Y 軸ボードのボードアドレス [i , k] と既に一致しており、且つ、宛先 Y 座標 j が上記第 j インタフェース変換 L S I : 3 2 - j のアドレス j と一致しているため、Y 軸クロスバースイッチ 3 1 でスイッチングする必要はない。このため、上記受信パケットは、上記 Y 軸ボード上の第 j インタフェース変換 L S I : 3 2 - j によって入出力線 2 0 0 [i , j , k] に出力される。

上記入出力線 2 0 0 [i , j , k] は、ボードアドレスとして Y、Z 座標値 [i , k] をもつ X 軸ボード 2 0 の X 軸クロスバースイッチ 2 1 に接続されている。上記 X 軸クロスバースイッチ 2 1 は、入出力配線 2 0 0 [i , j , k] からの上記パケットを受信すると、これを第 i コネクタ 2 2 - i が接続された第 i ポートに送出する。これによって、上記パケットは、パケットヘッダの宛先アドレス [i , j , k] に対応した演算ノードで受信処理されたことになる。

【 0 0 2 9 】

図 3 は、同一の Z 座標値 z をもつ X 軸クロスバースイッチ群と Y 軸クロスバースイッチ群との接続関係を示す。ここでは、図面の簡略化のために、各インタフ

エース変換 L S I に接続される光モジュールは省略してある。

L×M×N個の演算ノードからなる3次元クロスバーネットワークの各XY平面を構成する2次元クロスバーネットワークは、Y座標値と対応して用意されたM個のX軸ボード20-1～20-Mと、X座標値と対応して用意されたL個のY軸ボード30-1～30-Lとからなる。X軸ボード20-1～20-M上のM個のX軸クロスバースイッチ21-1～21-Mで形成されるL×M個の外部入出力ポートのうち、同一のX座標値iをもつM個の入出力ポートが、入出力線200[i, 1, z]～200[i, M, z]を介して、ボードアドレス[i, z]をもつY軸ボード30-i上のクロスバースイッチ31-iに接続される。

【0030】

Y軸ボード30-1～30-L上の各インタフェース変換LSIには、夫々に接続された入出力線200[i, 1, z]～200[i, M, z]と対応したアドレスを割り当てることができる。以下の説明では、Y軸ボード30-i上の配線200[i, j, z]と接続されたインタフェース変換LSIは、符号32[i, j, z]、上記インタフェース変換LSIの光モジュール33、34に接続された光信号入出力線（光ファイバ）は、符号300[i, j, z]で示すことにする。

【0031】

図4は、複数のXY平面を統合するZ軸ボード40とY軸ボード30との間の接続関係を示す。

図において、30-1-1～30-L-1は、Z座標値が「1」のY軸ボード群、30-1-2～30-L-2は、Z座標値が「2」のY軸ボード群を示し、30-1-N～30-L-Nは、Z座標値が「N」のY軸ボード群を示す。また、各Y軸ボードとZ軸ボードとの接続関係を光信号入出力線300[1, 1, 1]～300[L, M, N]によって示す。

光信号入出力線300に付されたアドレス値から明らかなように、3次元クロスバーネットワークの構成に必要なZ軸ボード40の枚数は、X、Y座標値の組み合わせ数に対応しており、各Z軸ボードは、N層のXY平面から引き出された同一X、Y座標値をもつ入出力線300間でパケット交換できるように、それぞ

れ特定のY軸ボード群と結合されている。

【0032】

図5は、Y軸ボード30上に搭載され、Y軸クロスバースイッチ31とZ軸クロスバースイッチ41との間でのパケット交換機能を備えるインタフェース変換LSI:32の1実施例を示す。

インタフェース変換LSI:32は、X軸クロスバースイッチ21と接続する入出力線200を収容するための第1入出力ポート51Aと、Y軸クロスバースイッチ31と接続する入出力線310を収容するための第2入出力ポート51Bと、光モジュール33と接続する入出力線320を収容するための第3入出力ポート51Cとを有し、上記各入出力ポートは入力ポートINと出力ポートOUTを備える。

CLK1は、X軸ボード20とY軸ボード30間の電気信号によるデータ転送のための基準クロックを示し、CLK2は、Y軸ボード30とZ軸ボード40間の光信号によるデータ転送のための基準クロックを示す。PLL回路60と61は、それぞれ上記クロックCLK1、CLK2に同期した内部クロックを発生する。

【0033】

第1入出力ポート51Aの入力ポートINで受信されたパケット信号は、位相調整回路52Aを経てデータバッファ53Aに入力される。上記位相調整回路52Aは、PLL回路60で生成される内部クロックに基づいて入力信号の位相を調整し、入力ポートINから入力されたパケット信号を、例えば、8バイト幅のデータとして上記データバッファ53Aに蓄える。

上記第1入出力ポート51Aの出力ポートOUTには、後述するように、セレクタ55Aで選択された第2、第3入出力ポートからの入力パケットが出力される。尚、上記セレクタ55Aの選択動作は、出力制御回路56Aによって制御されている。以下の説明では、入出力線200と対応する回路要素51A～56Aからなる回路部分を第1インタフェースユニットと呼び、入出力線310、320と対応する同様の回路部分を第2、第3インタフェースユニットと呼ぶことにする。

【 0 0 3 4 】

入力パケット制御回路 5 4 A は、位相調整回路 5 2 A の出力を監視し、パケットの先頭を示すスタートフラグを識別して、パケットヘッダの所定位置に設定された宛先 Y 座標を抽出する。上記入力パケット制御回路 5 4 A は、各パケットから抽出した宛先 Y 座標値と、予め記憶されている各インタフェース変換 L S I : 3 0 に固有の Y 座標値とを比較し、Y 座標値が一致した場合は、第 3 インタフェースユニットの出力制御回路 5 6 C に、また、Y 座標値が一致しない場合は、第 2 インタフェースユニットの出力制御回路 5 6 B に対して、それぞれパケットの出力許可要求信号 R E Q 1 を送り、出力制御回路 5 6 B または 5 6 C が返送するパケット出力許可信号 C 2 0 または C 3 0 に応答して、データバッファ 5 3 A 内のパケットを内部バス 6 2 A に出力する。上記内部バス 6 2 A は、第 2 インタフェースユニットの出力セレクタ 5 5 B と、第 3 インタフェースユニットの出力セレクタ 5 5 C に接続されている。

【 0 0 3 5 】

第 2 インタフェースユニットも、上記第 1 インタフェースユニットと同様に、位相調整回路 5 2 B、データバッファ 5 3 B、入力パケット制御回路 5 4 B、セレクタ 5 5 B、出力制御回路 5 6 B を備えている。

第 2 インタフェースユニットの入力ポート I N には、Y 軸クロスバースイッチ 3 1 によって Y 軸方向のスイッチングを終えたパケットが入力されるため、入力パケット制御回路 5 4 B では、入力パケットのヘッダから宛先 Z 座標値を抽出し、これを予め記憶されているボードアドレスの Z 座標値とを比較する。

Z 座標値が一致した場合は、第 1 インタフェースユニットの出力制御回路 5 6 A に、また、Z 座標値が一致しない場合は、第 2 インタフェースユニットの出力制御回路 5 6 B にパケットの出力許可要求信号 R E Q 2 を送り、出力制御回路 5 6 A または 5 6 C が返送するパケット出力許可信号 C 1 0 または C 3 0 に応答して、データバッファ 5 3 B 内のパケットを内部バス 6 2 B に出力する。上記内部バス 6 2 B は、第 1 インタフェースユニットの出力セレクタ 5 5 A と、第 3 インタフェースユニットの出力セレクタ 5 5 C に接続されている。

【 0 0 3 6 】

第 3 インタフェースユニットは、データバッファ 5 3 C、入力パケット制御回路 5 4 C、セクタ 5 5 C、出力制御回路 5 6 Cの他に、同期化回路 5 7、5 8 と、データ幅変換回路 5 9 を備えている。出力制御回路 5 6 Cは、要求信号 R E Q 1 または R E Q 2 を受信すると、要求発生順にパケット出力許可信号 C 3 0 を返送し、セクタ 5 5 C を切替えることによって、要求元と対応する内部バス 6 2 A または 6 2 B のパケットデータを同期化回路 5 7 に取り込む。

同期化回路 5 7 は、P L L 回路 6 1 が発生する内部クロックに同期して動作し、第 3 入出力ポート 5 1 C におけるデータ送信を外部クロック C L K 2 に同期させる。上記同期化回路 5 7 から出力された 8 バイト幅のパケットは、データ幅変換回路 5 9 によって、例えば、2 2 ビット幅のパケットデータに変換された後、第 1、第 2 入出力ポートにおけるデータ転送の 4 倍の速度（周波数）で出力ポート O U T から送出される。出力ポート O U T から送出されたパケットデータは、レーザ発光素子 3 3 T で光信号に変換され、光ファイバ 3 0 0 を介して Z 軸ボード 4 0 に送られる。尚、上記入出力ポート 5 1 C と、レーザ発光素子 3 3 T および受光素子 3 3 R を含む光モジュール 3 3 との間では、上記パケットデータの他に同期クロックも送受信される。

【 0 0 3 7 】

Z 軸クロスバースイッチ 4 1 で Z 軸方向のスイッチングを終えたパケットは、光ファイバ 3 0 0 を介して受光素子 3 3 R に入力され、電気信号に変換された後、上記第 3 入出力ポート 5 1 C の入力ポート I N に入力される。入力ポート I N で受信されたパケットデータは、データ幅変換回路 5 1 C で 8 バイト幅に変換された後、同期化回路 5 8 を経て、データバッファ 5 3 C に蓄積される。尚、同期化回路 5 8 は、P L L 回路 6 0 が発生する内部クロックに同期して動作している。

【 0 0 3 8 】

入力パケット制御回路 5 4 C は、同期化回路 5 8 からデータバッファ 5 3 C に供給されるパケットデータを監視し、パケットヘッダの宛先 Y 座標を抽出して、予め記憶されているインタフェース変換 L S I : 3 0 に固有の Y 座標値とを比較

する。Y座標値が一致した場合は、第1インタフェースユニットの出力制御回路56Aに、また、Y座標値が一致しない場合は、第2インタフェースユニットの出力制御回路56Bに対して、それぞれパケットの出力許可要求信号REQ3を送り、出力制御回路56Aまたは56Bが返送するパケット出力許可信号C10またはC20に応答して、データバッファ53Cから内部バス62Cにパケットデータを出力する。

内部バス62Cは、第1インタフェースユニットの出力セレクタ55Aと、第2インタフェースユニットの出力セレクタ55Bに接続されている。尚、第3入出力ポート51Cで入出力されるパケットは、既にY軸方向のスイッチングを終えた状態となっているため、実際の動作において、入力パケットが第2インタフェースユニットに転送されることはない。

【0039】

第1、第2インタフェースユニットにおける出力セレクタ55A、55Bも、セレクタ55Cと同様に、それぞれ出力制御回路56A、56Bによって制御され、内部バス62Bと62C、62Aと62Cの中からパケット出力を許可された内部バス選択する。

上記構成によって、上記インタフェース変換LSI:32は、第1、第2、第3の3つのインタフェースユニット間で、受信パケットの宛先アドレスと該LSIの固有アドレスとの関係に応じたパケット転送動作を行うことができ、これをY軸ボードに適用することによって、図1で説明した3次元配列の任意の演算ノード間のパケット転送が可能となる。

【0040】

上記インタフェース変換LSIでは、電気信号インタフェース側で使用する同期クロックCLK1と、光信号インタフェース側で使用する同期クロックCLK2とを独立させ、X軸クロスバースイッチとY軸クロスバースイッチとの間ではクロックCLK1に同期した電気信号によってパケット転送し、Y軸ボードとZ軸ボードとの間ではクロックCLK2に同期した光信号でパケット転送するようになっているため、位相調整回路52に必要な同期クロックCLK1の分配範囲を、例えば、同軸ケーブルで分配容易な範囲に局所化することができる。

【 0 0 4 1 】

従って、上記インタフェース変換 L S I を使用すれば、同一クロック源による同期クロック分配範囲、例えば、1 群の X、Y ボードで構成される L × M 個の演算ノードを含むクロスバーネットワークを単位として、クロック源の異なる複数のクロスバーネットワーク間を容易に相互接続できる。また、小規模クロスバーネットワークの需要には上記 L × M 個の演算ノードからなる基本ユニットで対処し、必要に応じて Z 軸ボードと光ファイバによる接続ユニット数を増加することによって、需要に応じたクロスバーネットワーク規模の拡張（ノード数の増設）が極めて容易になる。

【 0 0 4 2 】

図 6 は、インタフェース変換 L S I : 3 2 の実用的な形態の 1 例を示す。

ここに示した例では、図 5 で説明した回路構成を同一 L S I 基板上に 2 組搭載し、5 1 A - 1、5 1 B - 1、5 1 C - 1 を入出力ポートとする第 1 の回路部と、5 1 A - 2、5 1 B - 2、5 1 C - 2 を入出力ポートとする第 2 の回路部で、P L L 回路 6 0 と 6 1 を共用するようにしている。

このように、図 5 の回路構成を同一 L S I 基板上に複数組搭載し、Y 軸ボード 3 0 上でのインタフェース変換 L S I : 3 2 の占有面積を減らすことによって、3 次元クロスバーネットワークの小型化と、接続ノード数の増加が可能となる。

【 0 0 4 3 】

Z 軸ボード 4 0 に搭載されるインタフェース変換 L S I : 4 1 は、図 5 の回路構成において、入出力線 2 0 0 に接続される第 1 インタフェースユニットが不要となり、Z 軸クロスバースイッチに接続される第 2 インタフェースユニットと、光モジュール 3 3 に接続される第 3 インタフェースユニットのみが必要となる。この場合、内部バス上での出力パケットの競合が発生しないため、各インタフェースユニットは、入力パケット制御回路 5 4 B、5 4 C、セクタ 5 5 B、5 5 C、出力競合制御回路 5 6 B、5 6 C を省略した簡単な回路構成を採用できる。上記 Z 軸ボード用のインタフェース変換 L S I : 4 1 についても、図 6 と同様、同一 L S I 基板上に複数組の回路部を搭載した装置構成を採用できる。

【 0 0 4 4 】

図 7 は、Y 軸クロスバースイッチ 3 1 の主要部の構成を示す。

Y 軸クロスバースイッチ 3 1 は、入出力線 3 1 0 - j (j = 1 ~ M) と対応した複数のインタフェースユニット 7 0 - j をデータバスと制御線で相互接続した構成となっている。X 軸クロスバースイッチ 4 1 もこれと同様の構造をもつ。

【 0 0 4 5 】

各インタフェースユニット 7 0 - j は、入出力線 3 1 0 - j を接続するための入出力ポート 5 1 と、位相調整回路 5 2 と、データバッファ 5 3 と、入力パケット制御回路 5 4 と、セレクタ 5 5 と、出力制御回路 5 6、クロック C L K 1 に同期した内部クロックを発生する P L L 回路 6 0 とからなり、図 5 で説明したインタフェース変換 L S I 3 2 における第 1、第 2 のインタフェースユニットと類似した機能を備える。

即ち、入力パケット制御回路 5 4 は、位相調整回路からデータバッファ 5 3 に入力されるパケットデータを監視し、パケットヘッダから宛先 Y 座標を抽出すると、上記宛先 Y 座標値 q に対応した第 q インタフェースユニットの出力制御回路に接続された制御線 7 2 A - q に、パケット出力許可の要求信号 R E Q を発行する。この状態で、上記第 q インタフェースユニットから、制御線 7 3 A - q を介して出力許可信号 C N T を受信すると、入力パケット制御回路 5 4 は、データバッファ 5 3 からデータバス 7 1 にパケットデータを出力する。

【 0 0 4 6 】

上記データバス 7 1 は、各インタフェースユニットに専用のバスとなっており、セレクタ 5 5 には、インタフェースユニットの個数と対応した複数のバスが接続されている。上記セレクタ 5 5 を制御する出力制御回路 5 6 は、制御線 7 2 B と 7 3 B を介して、他の複数のインタフェースユニット 7 0 - j (j = 1 ~ M - 1) の入力パケット制御回路 5 4 に接続されている。

出力制御回路 5 6 は、何れかの制御線 7 2 B - k から要求信号 R E Q を受信すると、これに対応する制御線 7 3 B - k に出力許可信号 C N T を返送し、制御線 7 2 B - k と対応するバス 7 1 - k が選択されるようにセレクタ 5 5 を制御する。制御線 7 2 B から複数の要求信号 R E Q を受信した場合、出力制御回路 5 6 は

、所定のアルゴリズム、例えば、要求信号発生順に選ばれた制御線 7 3 B - k に対して出力許可信号 C N T を返送する。

【 0 0 4 7 】

図 8 は、Y 軸クロスバースイッチ 3 1 の実用的な 1 実施例を示す。

図 1 に示した Y 軸クロスバースイッチ 3 1 は、各インタフェース変換 L S I 3 2 との間で、例えば、8 バイト単位でパケットデータを送受信する。この場合、図 7 に示した各インタフェースユニット 7 0 - j の入出力ポート 5 1 に 8 バイト分の外部接続ピンが必要となる。また、各インタフェースユニット間の接続バス 7 1 に各々 8 バイト幅の信号線が必要となるため、L S I の構造が複雑化し、ポート数の増加が困難となる。

図 8 に示した実施例は、Y 軸クロスバースイッチ 3 1 を L S I 化した場合に必要となる外部接続ピン数を少なくし、構造を簡素化することによって L S I 化を容易にしたものであり、各入出力ポートが 2 バイト幅でパケットデータを送受信できるように、Y 軸クロスバースイッチ 3 1 を L S I : 3 1 A ~ 3 1 D に 4 分割 (バイトスライス構造) したことを特徴としている。

【 0 0 4 8 】

Y 軸クロスバースイッチ 3 1 に上記構造を採用した場合、図 5 に示した各インタフェース変換 L S I : 3 2 の第 2 入出力ポート 5 1 B において、入力ポート I N と出力ポート O U T に接続される入出力線 3 1 0 をそれぞれ 2 バイト幅をもつ第 1 ~ 第 4 の入出力線に分割し、図 8 に示すように、上記各 L S I : 3 1 A ~ 3 1 D に 2 バイト幅で並列的に入出力するようにしておく。また、各クロスバースイッチ 3 1 が 2 バイト幅のパケットを受信して経路選択できるようにするため、各パケットヘッダには、宛先ノードアドレスやデータ長等の情報を予め 4 重化して与えておく。上記分割構造は、X 軸クロスバースイッチ 4 1 も同様に適用できる。

【 0 0 4 9 】

多数の L S I と入出力線 (ケーブル) 接続用外部コネクタの搭載を必要とする Y 軸ボード 3 0 および Z 軸ボード 4 0 は、例えば、多層プリント配線基板の表面側に、クロスバースイッチ L S I : 3 1 (3 1 A ~ 3 1 D) と、入出力線 (例え

ば、同軸ケーブル) 200を接続するためのコネクタ群とを配置し、上記基板の裏面側に、インタフェース変換LSI群32と、それぞれ複数の光モジュール33を搭載した光ボード接続用のコネクタとを搭載し、各インタフェース変換LSI:32とクロスバースイッチLSI:31との接続(入出力線310)と、上記入出力線接続コネクタ群とインタフェース変換LSI群32との間の接続は、上記基板を貫通して一方の面から他方の面に延びて形成されるプリント配線によって達成する。

【0050】

Y軸ボードの構成に必要なその他の回路要素、例えば、上記クロスバースイッチLSI:31にデータ転送用クロック信号を分配するためのLSIは基板表面に、上記光モジュール33にクロック信号を分配するためのLSIは基板裏面に、上記クロスバースイッチLSI:31およびインタフェース変換LSI:32に初期値を設定するために必要なLSIと、終端抵抗やノイズ低減用のコンデンサは基板の両面に配置し、上記何れかの面に電源ボード搭載用コネクタを搭載すればよい。

【0051】

図9は、本発明による分散Exchanger方式の3次元クロスバーネットワークの第2の実施例を示す。

本実施例は、図1～図4で説明した3次元クロスバーネットワークを2組使用し、Y軸ボード30とZ軸ボード40の間に、図10で後述する4つのインタフェースユニット間のスイッチング機能を備えるスイッチングLSI:80を介在させることによって、第1クロスバーネットワークから第2クロスバーネットワークに、あるいはその逆方向にパケット転送を可能としたものである。

【0052】

図9において、符号Aが付されたX軸クロスバースイッチ21-1A～21A-M、Y軸クロスバースイッチ31A、Z軸クロスバースイッチ41A-1～41A-Mは、第1クロスバーネットワークを構成し、符号Bが付されたX軸クロスバースイッチ21-1B～21B-M、Y軸クロスバースイッチ31B、Z軸クロスバースイッチ41B-1～41B-Mは、第2クロスバーネットワークを

構成している。上記Y軸クロスバースイッチ31A、31Bは、それぞれ図4に示したY軸ボード30-1-1～30-L-Nのうちの1つに相当する。

【0053】

図9において、インタフェース変換LSI:42-1～42-Mは、1つのLSI基板に2ポート分の回路を含んでおり、実質的には、Z軸クロスバースイッチ41A-1～41A-M用のLSI群と、Z軸クロスバースイッチ41B-1～41B-M用のLSI群とに分かれている。

本実施例では、X軸クロスバースイッチ群に接続される各演算ノードに、ノードアドレスとして、3次元クロスバーネットワークにおける座標値 $[x, y, z]$ と、各ノードが所属する3次元クロスバーネットワークの識別子（セット識別子） $[t]$ を与えておき、演算ノードと対応したX軸クロスバースイッチの各入出力線200が $[x, y, z, t]$ のアドレスをもつようにする。また、各演算ノードから送信されるパケットのヘッダには、宛先アドレスとして、宛先ノードの座標値とネットワーク識別子を設定する。ここでは、第1、第2のクロスバーネットワークの識別子 t をそれぞれA、Bで表すことにする。

【0054】

Y軸クロスバースイッチ31A、31Bは、それぞれX軸クロスバースイッチに接続するための第1の入出力ポート群と、X軸クロスバースイッチに接続するための第2の入出力ポート群を備え、第1の入出力ポート群は、入出力線200によってX軸クロスバースイッチの入出力ポートに直結され、第2の入出力ポート群は、入出力線群340Aまたは340Bによって、スイッチングLSI:80-1～80-Mの第1または第2入出力ポートに接続される。

上記Y軸クロスバースイッチ31A、31Bでは、X軸クロスバースイッチから受信したパケットの宛先Y座標に従ってY軸方向のスイッチングを行い、Z軸方向またはネットワーク間スイッチングを要するパケットについては、スイッチングLSI側の出力ポートに送出し、これらのスイッチングを必要としないパケットについては、X軸クロスバースイッチ側の出力ポートに送出する。

【0055】

上記各スイッチングLSI:80-i ($i = 1 \sim M$) は、Y軸クロスバースイ

ッチ 3 1 A から受信したパケットの宛先アドレスをチェックし、宛先ネットワーク識別子 t が予め上記受信パケットの入力ポートに割り当てられたネットワーク識別子と異なる場合には、ネットワーク間のスイッチングを行ない、Z 軸方向のスイッチングを要するパケットは、光モジュール 8 3 側の出力ポートに送出し、Z 軸方向のスイッチングを必要としないパケットは、他方のネットワークに属した X 軸クロスバースイッチ側の出力ポートに送出する。上記光モジュール 8 3 は、光ファイバによって Z 軸ボード上の光モジュール 4 3 に接続されているため、前述した第 1 実施例と同様に、Z 軸クロスバースイッチ 4 1 A または 4 1 B による Z 軸方向のスイッチングが可能となる。

【 0 0 5 6 】

図 1 0 は、スイッチング L S I : 8 0 の構成を示す。

スイッチング L S I : 8 0 は、Y 軸クロスバースイッチ 3 1 との間でパケットを送受信する 2 つのインタフェースユニット 8 1 A、8 1 B と、光モジュール 8 3 に接続される 2 つのインタフェースユニット 8 1 C、8 1 D とからなり、これらのインタフェースユニット間でパケット交換を行う。図 9 に示したクロスバーネットワークに適用した場合、インタフェースユニット 8 1 A と 8 1 C が第 1 クロスバーネットワーク用、インタフェースユニット 8 1 B と 8 1 D が第 2 クロスバーネットワーク用となる。

インタフェースユニット 8 1 A と 8 1 B は、図 5 で説明したインタフェース変換 L S I : 3 2 の第 1 インタフェースユニットと同様の構成要素からなり、インタフェースユニット 8 1 C と 8 1 D は、上記インタフェース変換 L S I の第 3 インタフェースユニットと同様の構成要素からなっている。

【 0 0 5 7 】

スイッチング L S I : 8 0 を図 9 のクロスバーネットワークに適用する場合、インタフェースユニット 8 1 A の入力パケット制御回路 5 4 A で、入力ポート I N から入力された受信パケットのヘッダから宛先 Z 座標値 z と宛先ネットワーク識別子 t を抽出し、予め記憶している固有の Z 座標値およびネットワーク識別子（この例では A）と比較する。

宛先 Z 座標値 z と宛先ネットワーク識別子 t の双方が固有値と不一致の場合は

、第2クロスバーネットワーク用の光モジュール側のインタフェースユニット81Dの出力制御回路56Dにパケット出力許可要求を出し、出力許可信号を待って、データバッファ53Aから内部バス63Aにパケットを出力する。

宛先Z座標値zが一致し、宛先ネットワーク識別子t不一致の場合は、第2クロスバーネットワーク用のX軸クロスバースイッチ側のインタフェースユニット81Bの出力制御回路56Bに、宛先Z座標値zが不一致で、宛先ネットワーク識別子tが一致した場合は、第1クロスバーネットワーク用の光モジュール側のインタフェースユニット81Cの出力制御回路56Cにパケット出力許可要求を出し、出力許可信号を待って、データバッファ53Aから内部バス63Aにパケットを出力する。

【0058】

インタフェースユニット81Bの入力パケット制御回路54Bは、固有値として入力パケット制御回路54Aとは異なるネットワーク識別子（この例ではB）を記憶している。入力パケット制御回路54Bは、受信パケットの宛先Z座標値zと宛先ネットワーク識別子tの双方が固有値と不一致の場合は、第1クロスバーネットワーク用の光モジュール側のインタフェースユニット81Cの出力制御回路56Cに、宛先Z座標値zが一致し、宛先ネットワーク識別子tが不一致の場合は、第1クロスバーネットワーク用のX軸クロスバースイッチ側のインタフェースユニット81Aの出力制御回路56Aに、宛先Z座標値zが不一致で、宛先ネットワーク識別子tが一致した場合は、第2クロスバーネットワーク用の光モジュール側のインタフェースユニット81Dの出力制御回路56Dに、パケット出力許可要求を出す。そして、出力制御回路56A、56Cまたは56Dからの出力許可信号を待って、データバッファ53Bから内部バス63Bにパケットを出力する。

【0059】

インタフェースユニット81Cと81Dの入力ポートINには、既にY軸、Z軸方向およびネットワーク間のスイッチングを完了したパケットが入力されるため、受信パケットをそれぞれY軸クロスバースイッチ側のインタフェースユニット81A、81Bに転送すればよい。従って、インタフェースユニット81Cと

8 1 D の入力パケット制御回路 5 4 C、5 4 D は、入力ポート I N からパケットが入力された時、それぞれ転送先インタフェースユニットの出力制御回路 5 6 A、5 6 B にパケット出力許可要求を出し、出力許可信号を待って、データバッファ 5 3 C、5 3 D から内部バス 6 3 C、6 3 D にパケットを出力すればよい。しかしながら、図 1 0 に示した実施例では、スイッチング L S I : 8 0 の用途を汎用化するため、内部バス 6 3 C と 6 3 D をインタフェースユニット 8 1 A、8 1 B の双方の出力セレクタ 5 5 A、5 5 B に接続し、パケット転送先を選択できるようにしてある。

【 0 0 6 0 】

インタフェースユニット 8 1 A、8 1 B では、それぞれのセレクタ 5 5 A、5 5 B に他の 3 つのインタフェースユニットの内部バスを接続し、そのうちの 1 つを出力制御回路 5 6 A、5 6 B で選択することによって、3 つの入力ポートからの受信パケットを出力ポート O U T に選択的に送出できるようになっている。また、インタフェースユニット 8 1 C、8 1 D では、それぞれのセレクタ 5 5 C、5 5 D に Y 軸クロスバースイッチ側の 2 つの内部バス 6 3 A、6 3 B を接続し、そのうちの 1 つを出力制御回路 5 6 C、5 6 D で選択することによって、2 つの入力ポートからの受信パケットを出力ポート O U T に選択的に送出できるようになっている。

【 0 0 6 1 】

上記第 2 実施例の構成によれば、スイッチング L S I : 8 0 を 4 次元目のスイッチとして利用し、Z 軸クロスバースイッチを 2 組の 3 次元クロスバーネットワークで共用させることによって、ハードウェア規模を倍増することなく、演算ノード数の倍増することができる。また、上述した分散 Exchanger 方式クロスバースイッチ機能を併せ持つスイッチング L S I : 8 0 を Z 軸クロスバースイッチ側にも適用すれば、安価な間接続コストで更に次元数の高いクロスバーネットワークを構成することが可能となる。

【 0 0 6 2 】

図 1 1 は、本発明による分散 Exchanger 方式の 3 次元クロスバーネットワークの第 3 の実施例を示す。

本実施例は、図 1 ～ 図 4 で説明した 3 次元クロスバーネットワークにおける Z 軸クロスバースイッチ 4 1 の機能をスイッチング L S I : 8 0 によって実現したことを特徴とする。

図において、Y 軸クロスバースイッチ 3 1 A、3 1 B、3 1 C、3 1 D は、それぞれ異なる Z 座標値 z 、例えば、 $z = 1, 2, 3, 4$ をもつ。

【 0 0 6 3 】

本実施例では、第 2 実施例と同様に、Y 軸クロスバースイッチ 3 1 A と 3 1 B の同一 Y 座標をもつ入出力ポート同士を第 1 のスイッチング L S I 群 8 0 - 1 ～ 8 0 - M により接続し、これらの L S I 群におけるパケットのスイッチング機能を利用して、Y 軸クロスバースイッチ 3 1 A と 3 1 B の間での Z 軸方向のスイッチングを行う。これと同様に、Y 軸クロスバースイッチ 3 1 C と 3 1 D を第 2 のスイッチング L S I 群 8 2 - 1 ～ 8 2 - M により接続し、これらの L S I 群によって、Y 軸クロスバースイッチ 3 1 C と 3 1 D の間の Z 軸方向のスイッチングを行う。

上記第 1、第 2 のスイッチング L S I 群は、同一 Y 座標をもつ L S I : 8 0 - i と 8 2 - i ($i = 1 \sim M$) の間を光モジュール 8 3、8 4 と光ファイバ 3 0 0 を介して結合する。

【 0 0 6 4 】

第 1 L S I 群に属した各スイッチング L S I : 8 0 - i のインタフェースユニット 8 1 A、8 1 B では、それぞれの入力パケット制御回路によって、Y 軸クロスバースイッチ 3 1 A または 3 1 B から受信したパケットの宛先 Z 座標値 z に応じて、 $z = 1$ または $z = 2$ の場合はインタフェースユニット 8 1 A と 8 1 B の間で、 $z = 3$ の場合はインタフェースユニット 8 1 C 宛に、 $z = 4$ の場合はインタフェースユニット 8 1 D 宛にパケットを転送制御する。一方、第 2 L S I 群に属した各スイッチング L S I : 8 2 - i のインタフェースユニット 8 1 A、8 1 B では、Y 軸クロスバースイッチ 3 1 C または 3 1 D から受信したパケットの宛先 Z 座標値 z に応じて、 $z = 1$ の場合はインタフェースユニット 8 1 C 宛に、 $z = 2$ の場合はインタフェースユニット 8 1 D 宛に、 $z = 3$ または $z = 4$ の場合はインタフェースユニット 8 1 A、8 1 B 間でパケットを転送制御する。

本実施例によれば、第 1、第 2 のスイッチング L S I 群 8 0、8 2 のパケットスイッチング機能を利用して、ノード数が $L \times M \times 4$ の分散 Exchanger 方式 3 次元クロスバーネットワークを構成できる。

【 0 0 6 5 】

【発明の効果】

以上の説明から明かなように、本発明によれば、多次元クロスバーネットワークおよび並列計算機システムの少なくとも一部のクロスバースイッチ間に光インタフェースを備えたスイッチング装置を介在させることによって、システム性能を低下させることなく、接続可能なノード数を増加できるという効果がある。

【図面の簡単な説明】

【図 1】

本発明による Exchanger 方式の 3 次元クロスバーネットワークにおける X、Y、Z 軸クロスバースイッチ間の接続関係を説明するための図。

【図 2】

演算ノードが搭載されるノードボード 1 0 の構成を示す図。

【図 3】

本発明の分散 Exchanger 方式の 3 次元クロスバーネットワークにおける X 軸クロスバースイッチと Y 軸クロスバースイッチとの接続関係を示す図。

【図 4】

本発明の分散 Exchanger 方式の 3 次元クロスバーネットワークにおける Y 軸クロスバースイッチと Z 軸クロスバースイッチとの接続関係を示す図。

【図 5】

インターフェース変換 L S I : 3 2 の構成を示すブロック図。

【図 6】

インターフェース変換 L S I : 3 2 の実用的な形態の 1 例を示す図。

【図 7】

Y 軸クロスバースイッチの主要部を示す図。

【図 8】

Y 軸クロスバースイッチの変形実施例を示す図。

【図 9】

本発明によるExchanger方式の3次元クロスバーネットワークの第2の実施例を示す図。

【図 1 0】

図9におけるスイッチングLSI：80の1実施例を示す図。

【図 1 1】

本発明によるExchanger方式の3次元クロスバーネットワークの第3の実施例を示す図。

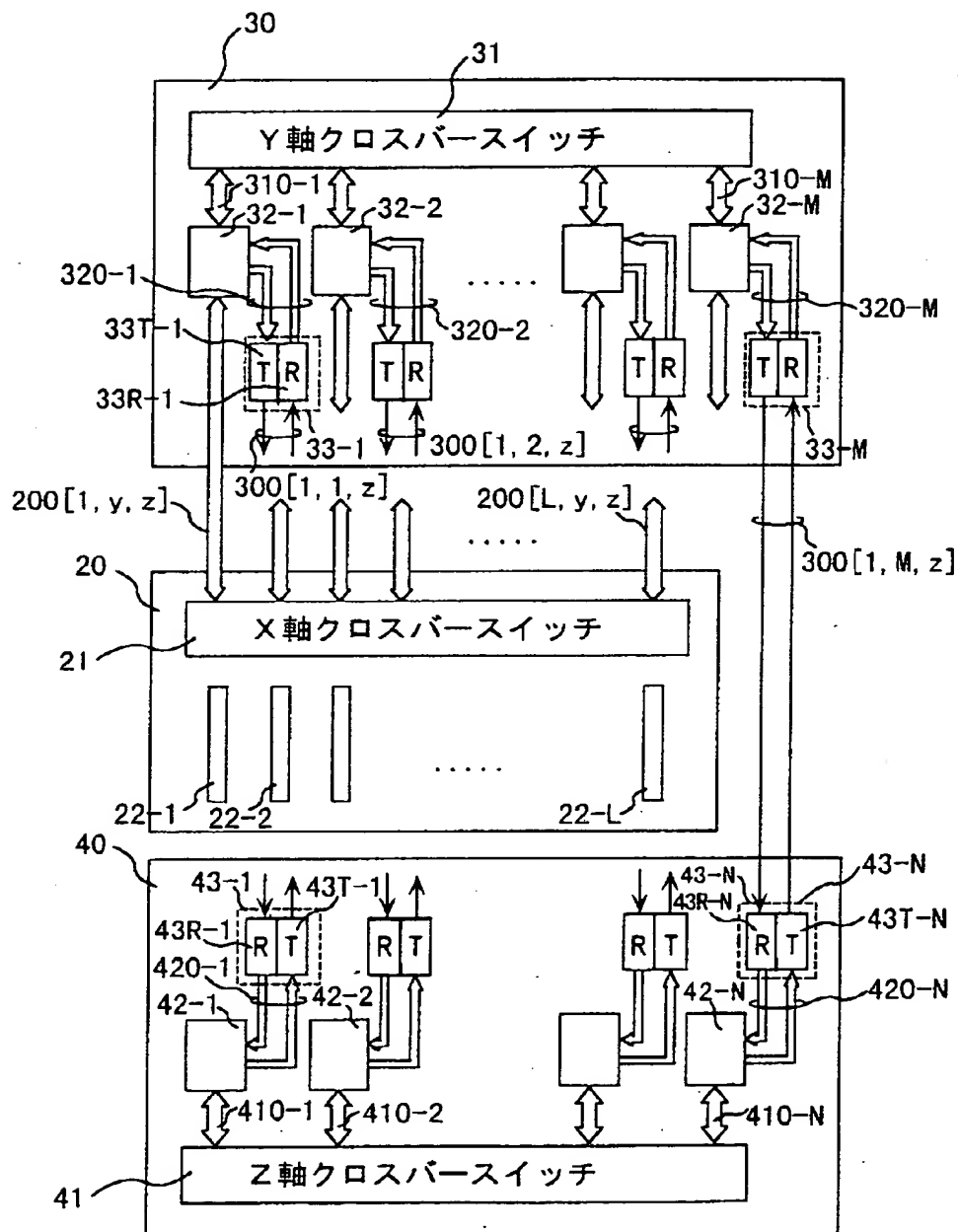
【符号の説明】

10：ノードボード、 20：X軸ボード、
21：X軸クロスバースイッチ、 22：ノードボード接続コネクタ、
30：Y軸ボード、 31：Y軸クロスバースイッチ、
32：インタフェース変換LSI、 33：光モジュール、
40：Z軸ボード、 42：インタフェース変換LSI、 43：光モジュール、
51：入出力ポート、 52：位相調整回路、 53：データバッファ、
54：入力パケット制御回路、 55：セレクタ、 56：出力制御回路、
57、58：同期化回路、 59：データ幅変換回路、 60、61：PLL回路、
200：入出力線（同軸ケーブル）、 300：入出力線（光ファイバ）。

【書類名】 図面

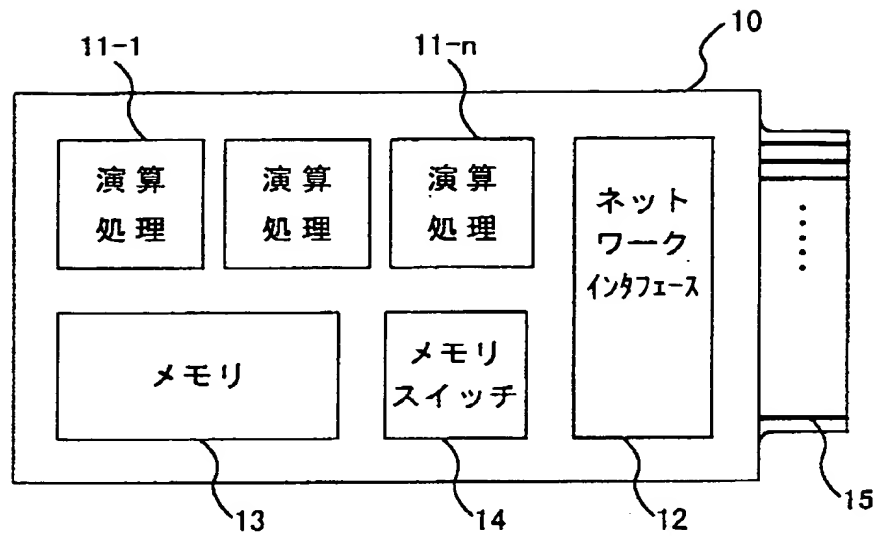
【図 1】

図 1



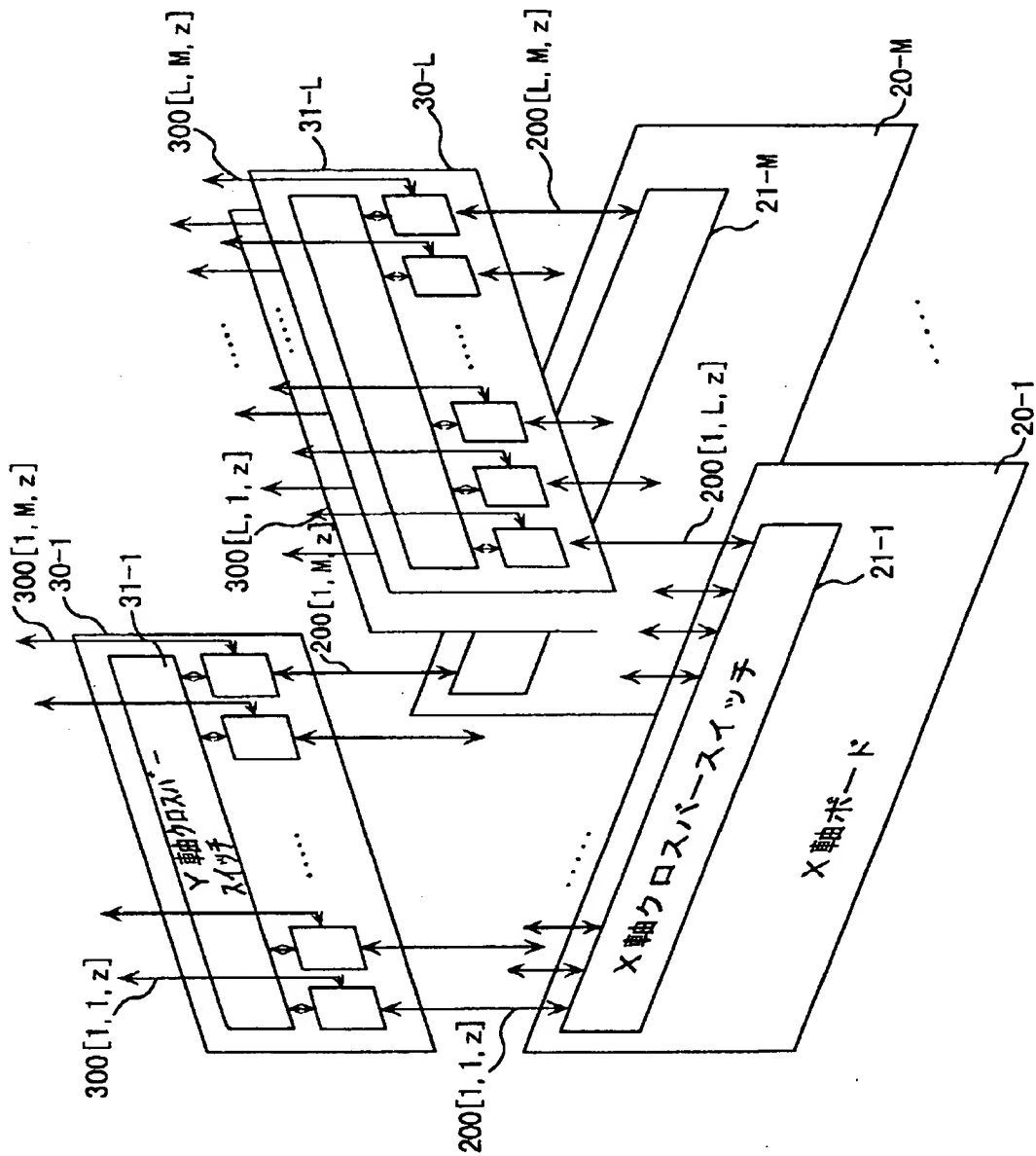
【図 2】

図 2



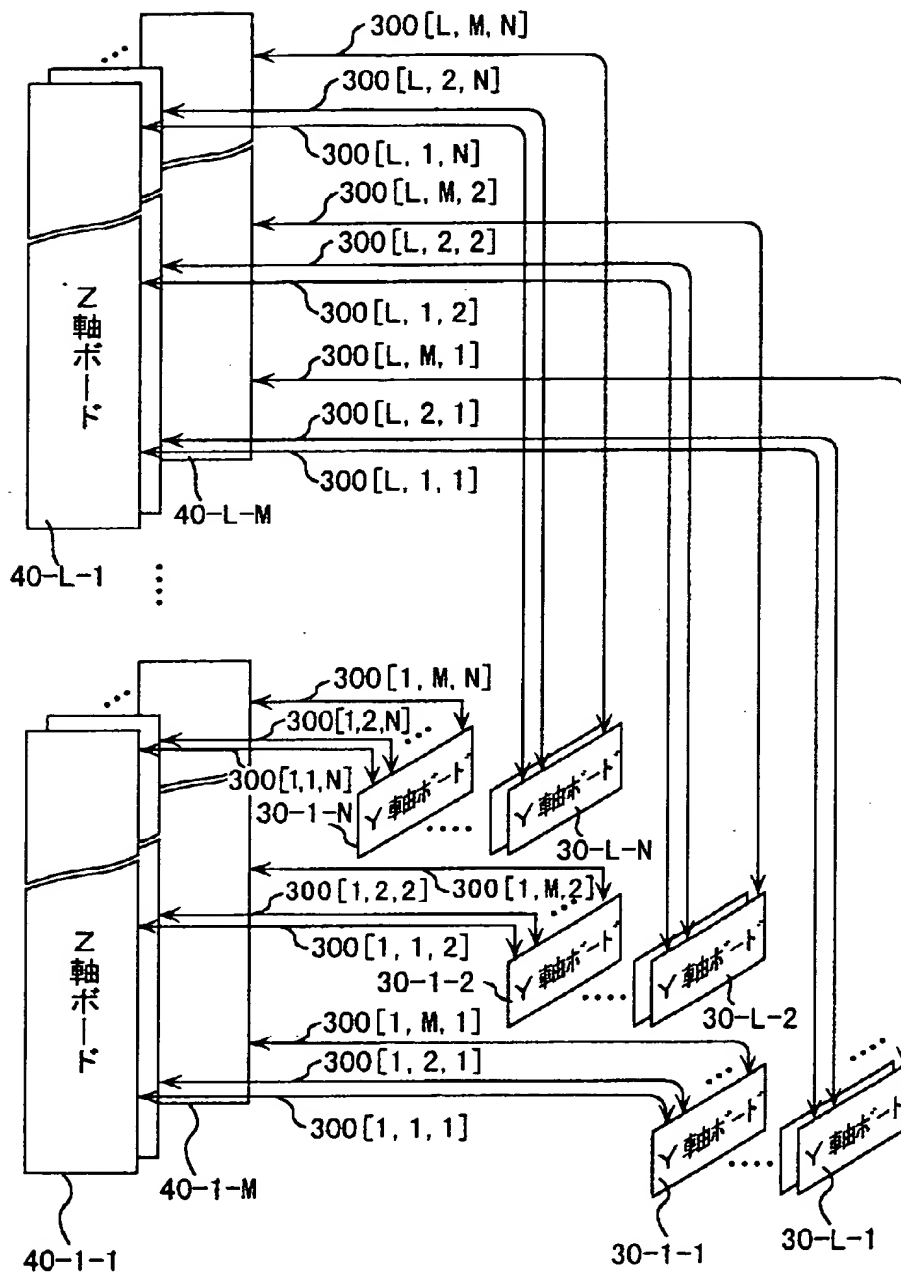
【図 3】

図 3



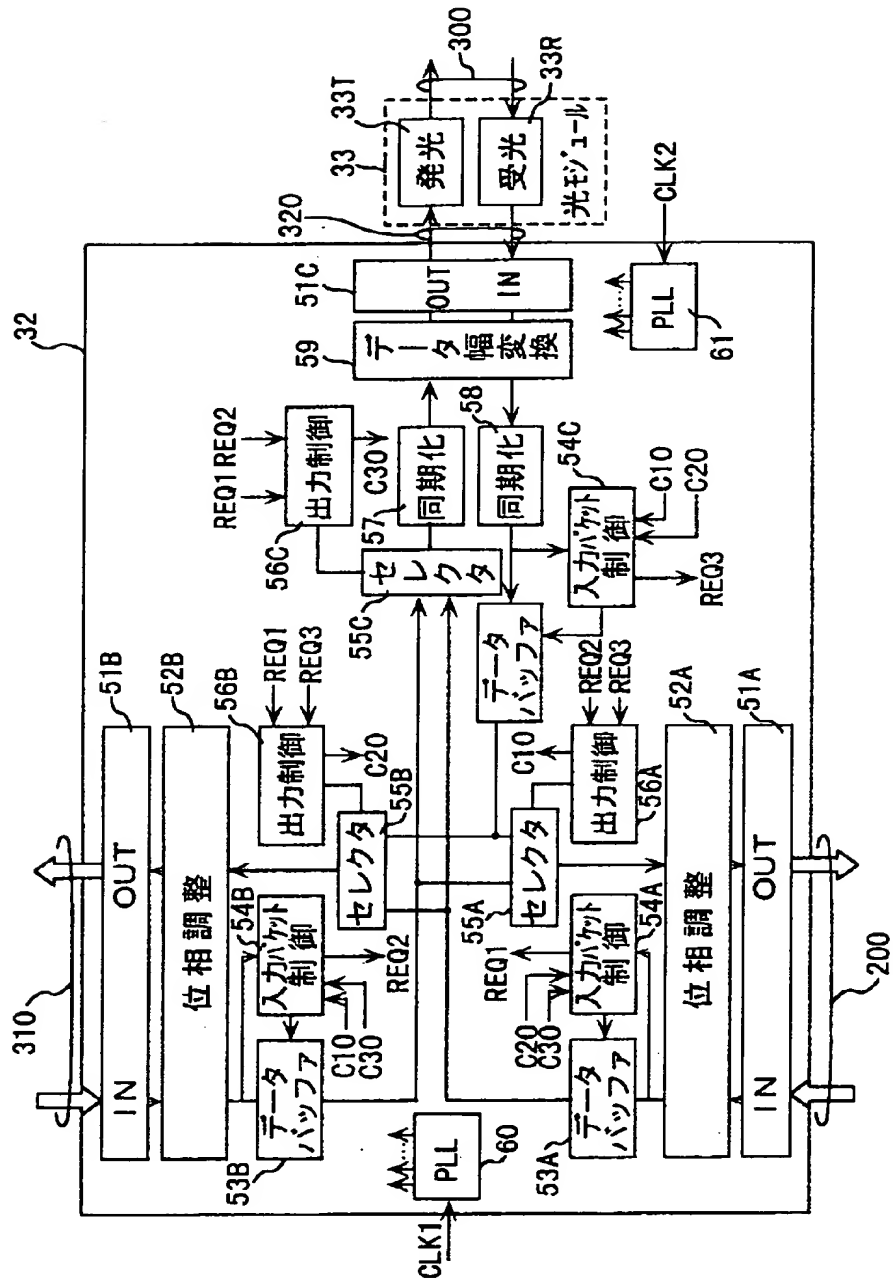
【図 4】

図 4



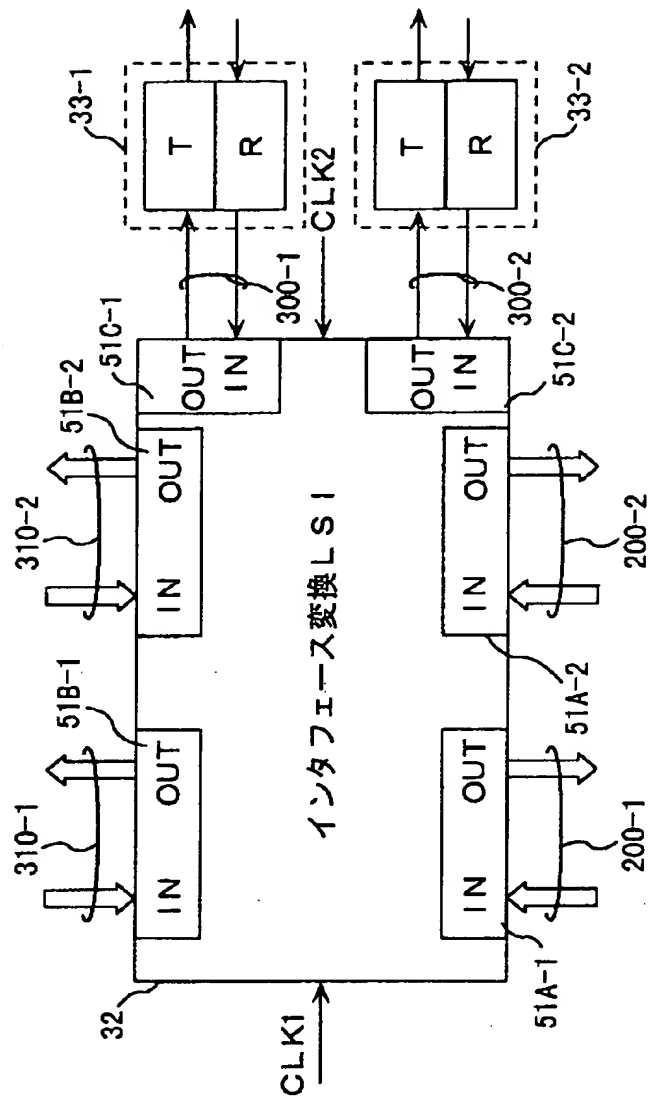
【图 5】

图 5



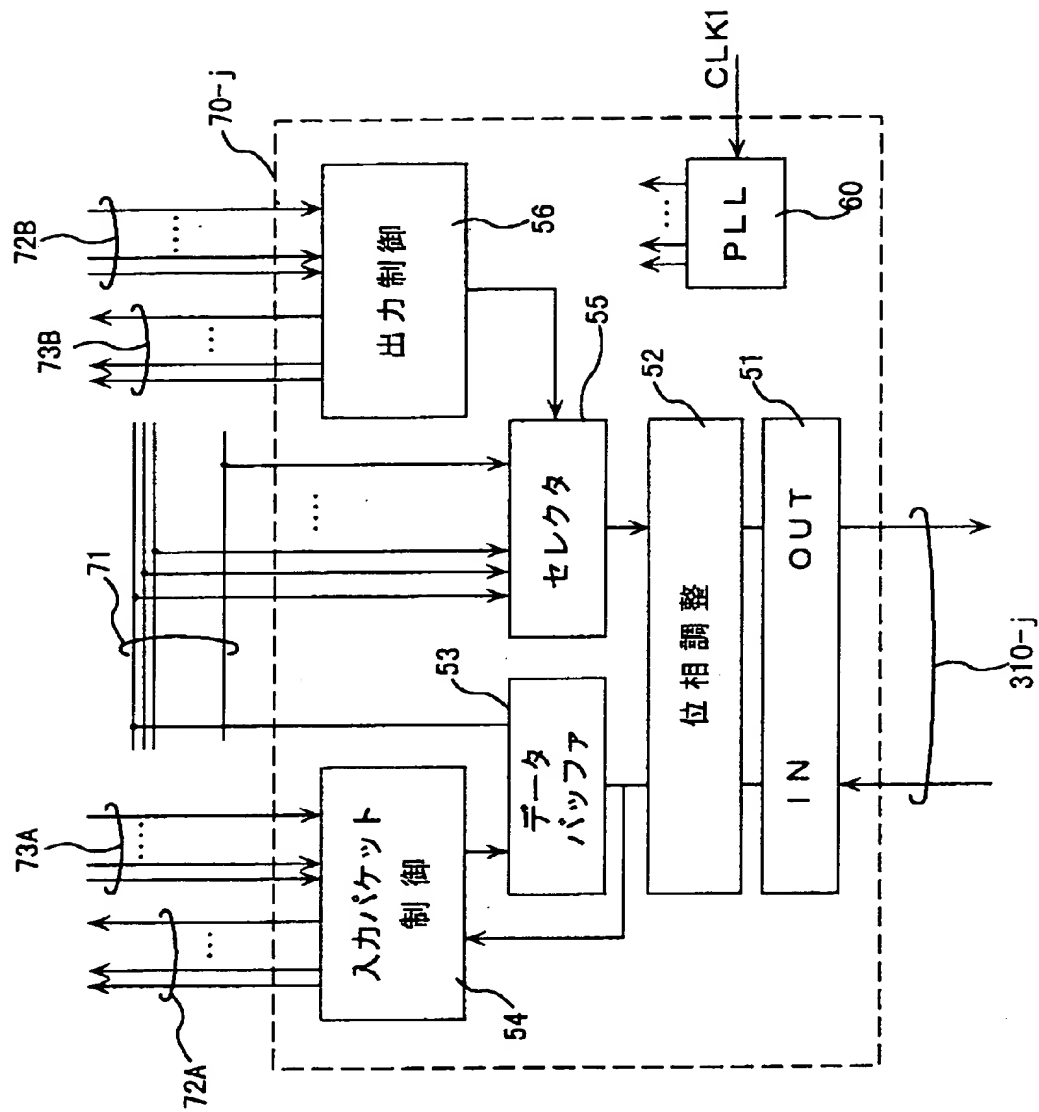
【図 6】

図 6



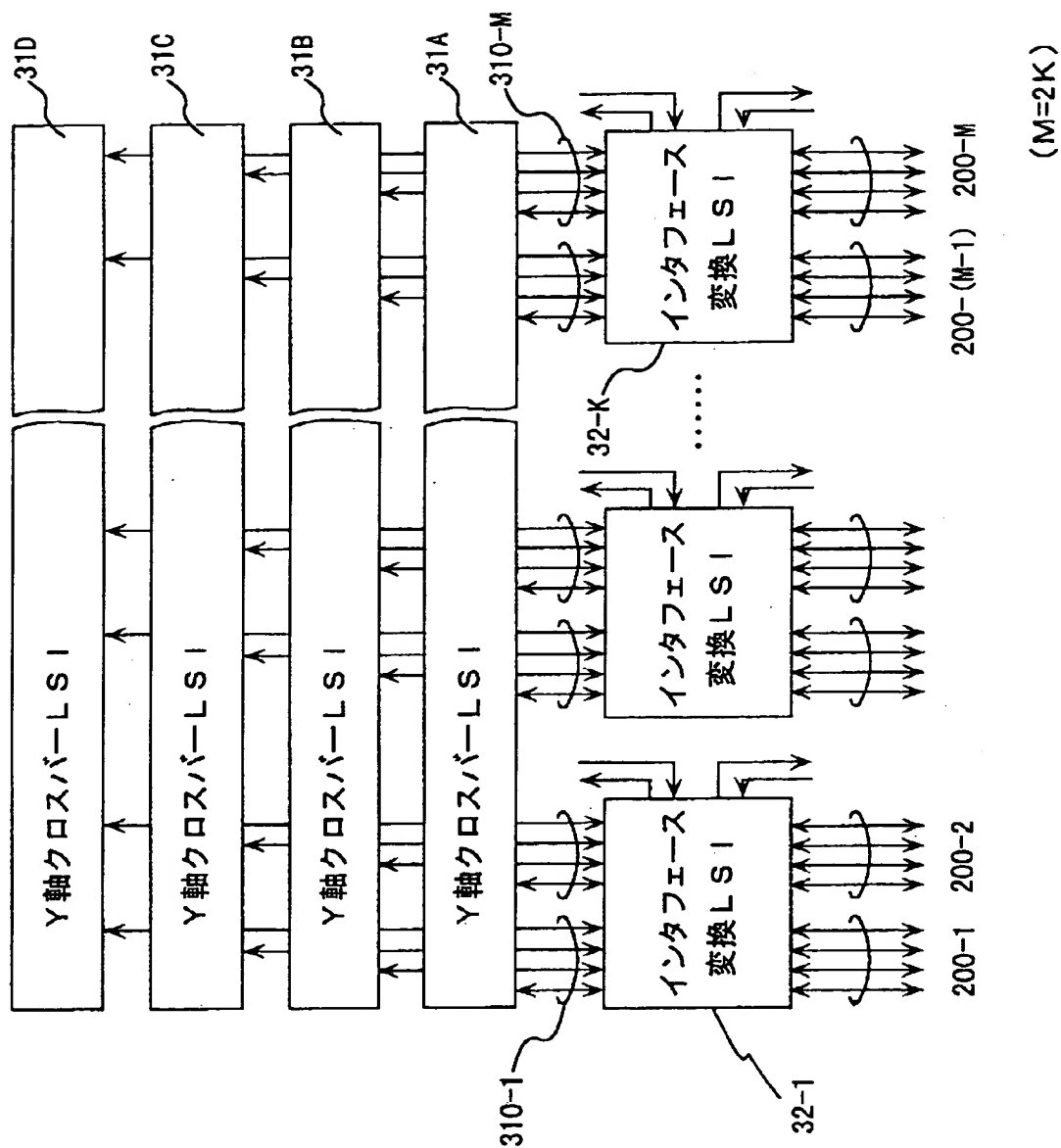
【図 7】

図 7



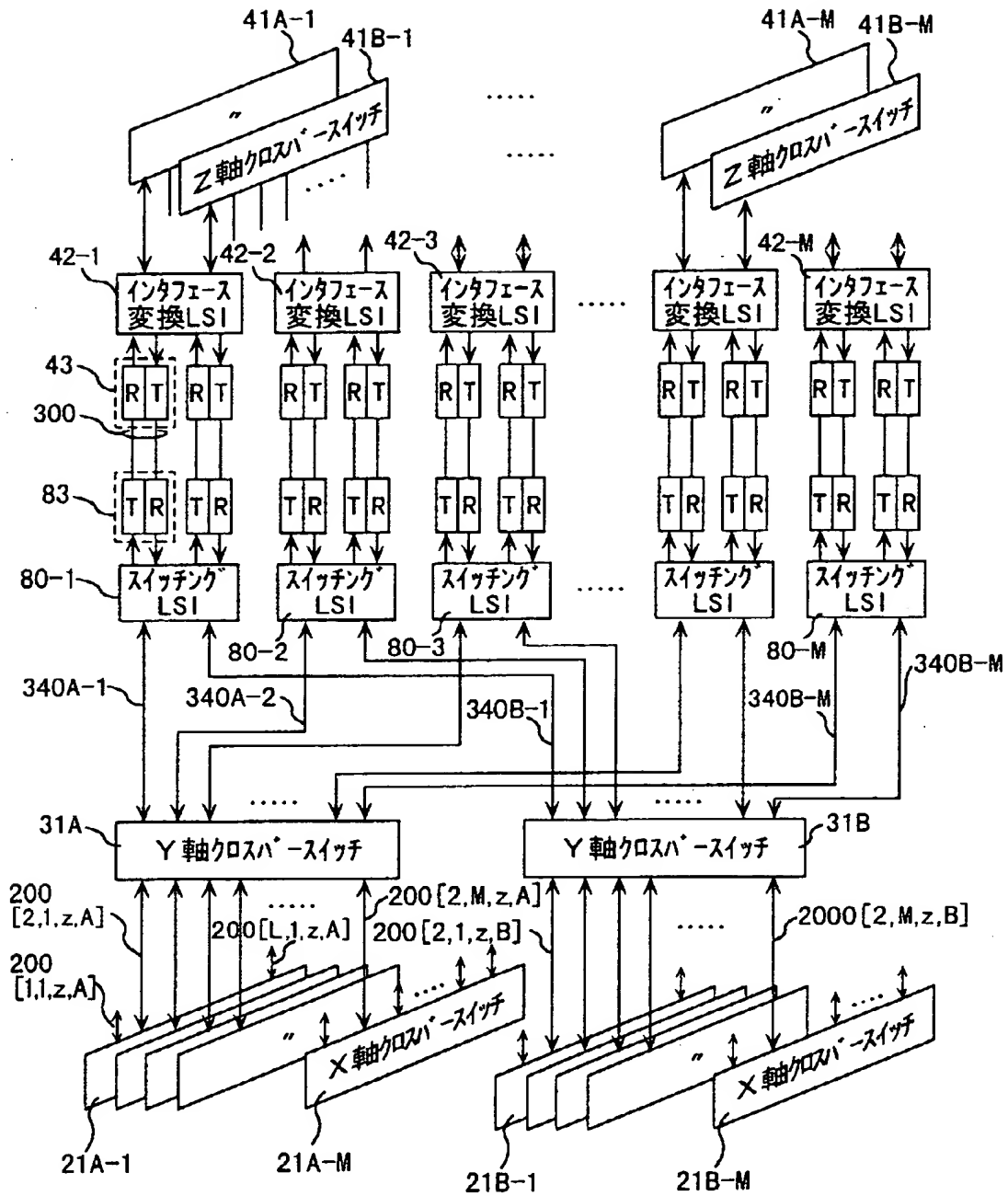
【図 8】

図 8



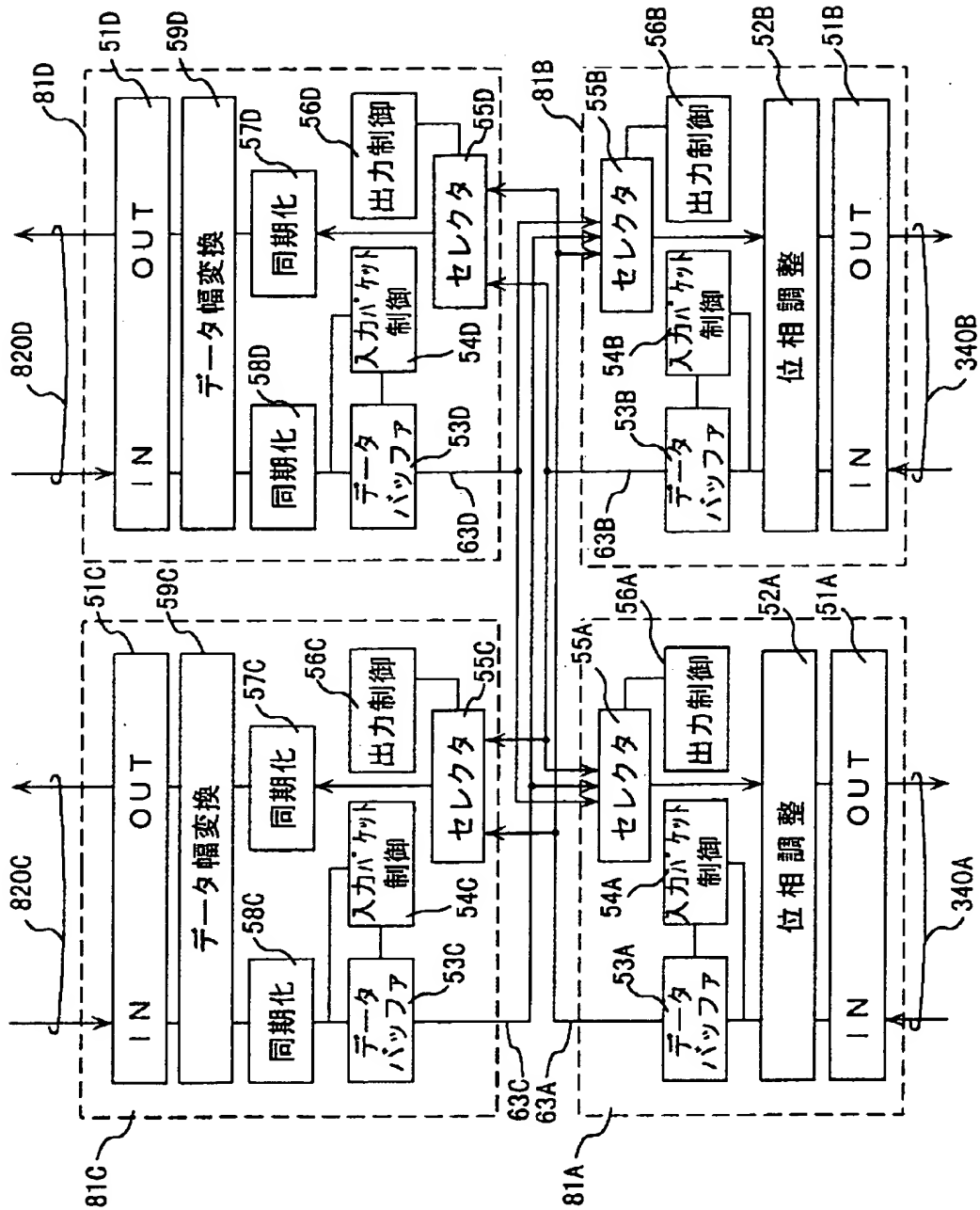
【図 9】

図 9



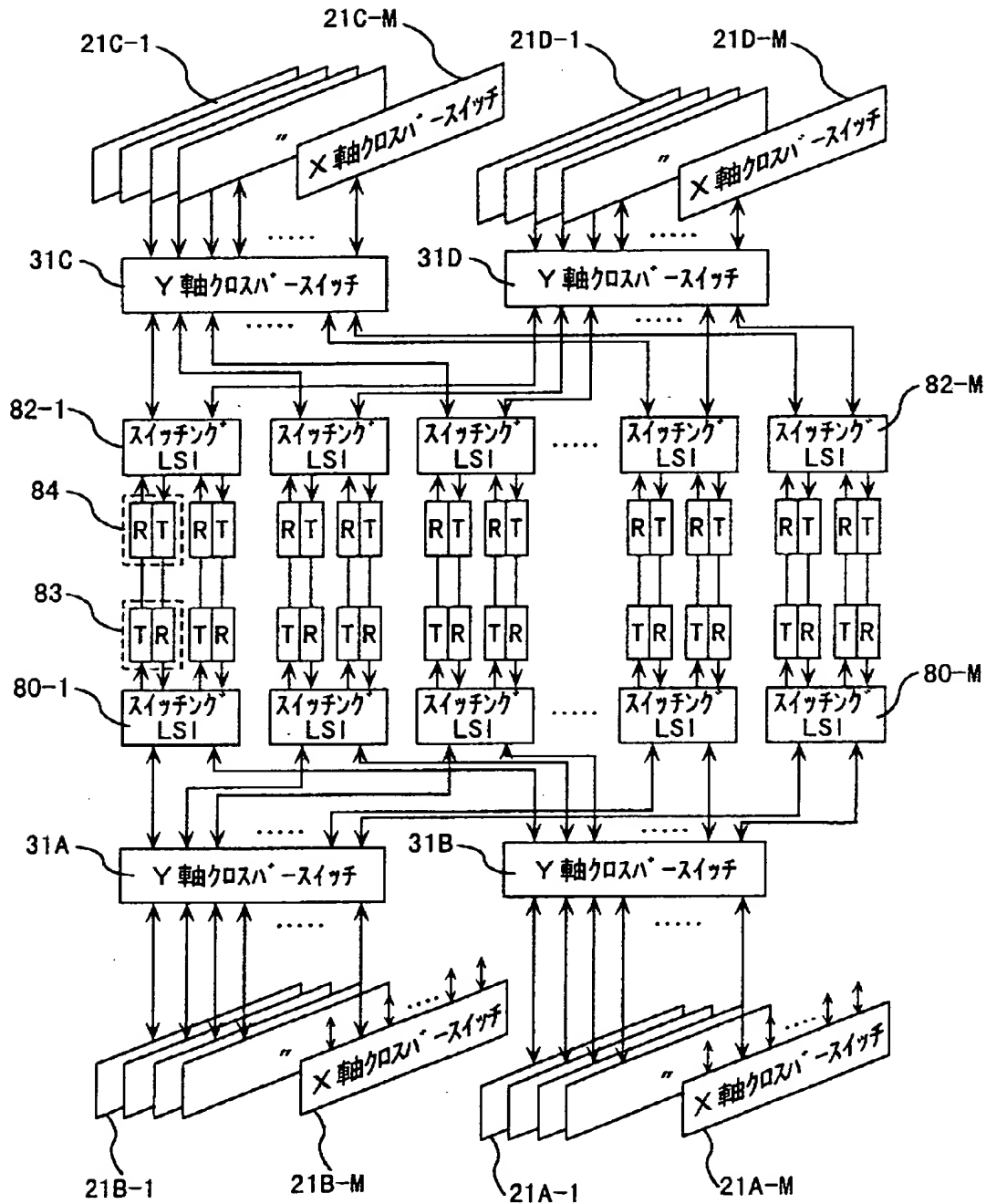
【図10】

図 10



【図 11】

図 11



【書類名】 要約書

【要約】

【課題】 分散Exchanger方式の多次元クロスバーネットワークおよび並列計算機システムにおいて、接続可能なノード数を増加すること。

【解決手段】 $L \times M \times N$ 個の演算装置を接続する分散Exchanger方式の3次元クロスバーネットワークにおいて、X軸クロスバースイッチ21とY軸クロスバースイッチ31との間に、光モジュール33を備えたインタフェース変換LSI32を介在させ、Z軸クロスバースイッチ41へのパケット転送を上記インタフェース変換LSIで行い、Y軸、Z軸クロスバースイッチ間のパケット転送路に光ファイバ300を適用する。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所